Complete chloroplast genomes of three copal trees (Bursera: Bullockia): comparative analysis and phylogenetic relationships

Gustavo P. Lorenzana
Northern Arizona University
Yessica Rico
yessica.rico@inecol.mx

INECOL, Mexico

Research Article

Keywords: Bursera, plastome genome, assembly and functional annotation, chloroplast structure, phylogenetics, Mexico

Posted Date: December 21st, 2023

DOI: https://doi.org/10.21203/rs.3.rs-3749786/v1

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Molecular Biology Reports on March 9th, 2024. See the published version at https://doi.org/10.1007/s11033-024-09304-z.

Abstract

Background

*Bursera*trees are conspicuous elements of the tropical dry forests in the Neotropics, which have a significant cultural value due to their fragrant resins (incense), wood sources (handcrafts), and ecological benefits. Despite its relevance, genetic resources developed for the genus are scarce.

Methods and results

We sequenced and analyzed the complete chloroplast genome structure and functional annotation of three *Bursera* species of the *Bullockia* section: *Bursera cuneata*, *B. palmeri*, and *B. bipinnata*. The chloroplast (Cp) genome sizes ranged from 159,824 to 159,872 bp in length, including a large single-copy (LSC) region from 87,668 to 87,656 bp, a small single-copy (SSC) from 18,581 to 18,571 bp, and two inverted repeats regions (IRa and IRb) of 26,814 bp each. The three Cp genomes consisted of 135 genes, of which 90 were functional, 37 tRNAs, and 8 rRNAs. The Cp genomes were relatively conserved, with the LSC region exhibiting the greatest nucleotide divergence (*psbJ, trnQ-UCC, trnG-UCC,* and *petL* genes), whereas few changes were observed in the IR border regions. Between 589 to 591 simple sequence repeats were identified. Phylogenetic relationships within Burseraceae for each Cp region (LSC, SSC, IRa, and IRb) using Bayesian inference confirmed that *Commiphora* is the sister taxa of *Bursera*. Only the phylogenetic trees based on the SSC and LSC regions resolved the close relationship between *B. bipinnata* and *B. palmeri* within *Bursera*.

Conclusion

Our work contributes to the development of *Bursera*'s genomic resources for taxonomic, evolutionary, and ecological-genetic studies.

Introduction

Bursera (Jacq. ex L.) is a diversified genus of small to medium-sized trees native to the Neotropics, and it is Mexico the center of biodiversity and endemicity with about 90 species [1, 2]. *Bursera* trees produce aromatic compounds, mostly terpenoids, which are released as resins as a deterrent response to herbivores [3, 4]. These fragrant resins have been used as incense (copal) in religious and medicinal practices since the Mesoamerican civilizations [5]. In rural areas, *Bursera* trees are sources of wood for the elaboration of handcrafts (e.g., alebrijes, masks), also act as living barriers against the wind, and can be used for ecological restoration [6]. They are conspicuous, at times dominant, elements of the Tropical Dry Forests (TDFs) [7, 8], providing food resources and shelter for a wide range of animals, primarily frugivore birds that consume the seeds of the plant [9, 10], and insects that pollinate it [11].

Bursera is divided into two sections, *Bullockia* and *Bursera*, based on differences in reproductive structures and bark morphology [1, 12]. The closest relatives of *Bursera* are species of the myrrh genus,

Commiphora, which migrated from North America to northern Africa and central Asia in the Early Eocene [12], followed by the divergence of the sections *Bursera* and *Bullockia* (approx. 40 million years ago (Ma)) [13]. The diversification of the two lineages in *Bursera* was in parallel with the aridification and expansion of the TDFs in north and central Mexico during the Miocene (20 to 17 Ma) [13].

Most genetic studies in *Bursera* are on its phylogenetic relationships based on a few nuclear (internal and external transcribed spacers, ribosomal regions, the phosphoenolpyruvate carboxylase, and the nitrate reductase) and chloroplast genes (*psbA-trnH* and *rps16*) [12–16] Recently, Rico et al. [17] reported the first nuclear whole-genome sequences of three species, commonly called copales, *Bursera bipinnata* (Moc. & Sessé ex DC.) Engl., *Bursera cuneata* (Schltdl.) Engl., and *Bursera palmeri* S.Watson, whereas Burseraceae chloroplast genomes have only been sequenced for Old World species (e.g.,[18, 19]).

Having its own DNA, the chloroplast has paramount functional importance as it is the specialized organelle where photosynthesis takes place. In angiosperms, chloroplast (Cp) genome size ranges from 107–218 kb and is generally arranged in a circular DNA molecule that displays a highly conserved tetragonal structure, consisting of a long (LSC) and a short single-copy (SSC) regions, interspersed by two inverted repeats (IR) regions [20]. Most of the Cp genome sequence encodes protein-coding genes, transfer RNA (tRNA), and ribosomal RNA (rRNA). The Cp genome sequencing is widely used in plant comparative genomics and phylogenetic relationships, while also facilitating the study of intraspecific genetic diversity and gene flow through the development of highly polymorphic markers such as Simple Sequence Repeats (SSRs) [21].

Here, we analyzed the Cp genomes of three *Bursera* species of the *Bullockia* section endemic to Mexico: *B. bippinata, B. cuneata*, and *B. palmeri.* Specifically, we aimed to (*i*) compare the structure and functional annotation of the complete Cp genomes; (*ii*) identify regions of high divergence and subject to selection; (*iii*) characterize the SSRs composition, and (*iv*) reconstruct the phylogenetic relationships by including the Cp genomes of other Burseraceae species.

Material and methods

Plant material, DNA extraction, and sequencing

Bursera bipinnata, *B. cuneata*, and *B. palmeri*, are deciduous trees with no exfoliating barks that can reach a height of 8–10 meters. Of the three species, *B. bipinnata* is the species with the widest distribution along the TDF in the Pacific coast of Mexico and is the species most used for the extraction of aromatic resins. *Bursera palmeri* occurs in fragments of TDF in the Central Mexican Plateau and *B. cuneata* has a smaller distribution across central south Mexico and is used in Michoacán for elaborating religious figures and masks [8]. Also, anti-inflammatory and antihistaminic compounds have been isolated from *B. cuneata* [22]. For each species, we collected leaf tissue samples from a single individual: *B. bipinnata* in Jalisco, *B. palmeri* in Querétaro, and *B. cuneata* in Ciudad de México (sampling permit

SEMARNAT SGPA/DGGFS/712/1062/18). We preserved the leaves in sealed plastic bags containing silica gel until DNA was extracted.

We extracted genomic DNA from 20 mg of dried tissue using the CTAB technique with pre-washing steps to eliminate excess of polyphenols [23]. The yield was estimated using a Qubit fluorimeter (Thermo Fisher Scientific, Waltham, MA, USA), and sequencing libraries were prepared with Illumina Truseq Nano DNA Library Prep kits (Illumina, San Diego, CA, USA) using 100 ng of high-molecular-weight DNA. Libraries were sequenced in pair-end mode with a read length of 150 bp using an Illumina NovaSeq S4 platform to obtain ~ 140–170 million reads (see [17]for more details). Genome sequencing was performed by NC State Genomic Sciences Laboratory (North Carolina, USA). Read cleaning and QC were performed with fastp [24].

Chloroplast genome assembly and annotation

The Cp genome assembly was performed with GetOrganelle [25], followed by visual inspection and manual curation using UGENE toolkit [26], which was performed on a multiple alignment of the three *Bursera* sequences against publicly available Cp sequences of *Commiphora* and *Boswellia* [18, 19]. Annotation was automatized using the CPGAVAS2 web platform [27], and circular plots of the Cp genomes were obtained with Chloroplot [28]. Finally, we prepared GenBank files for submission to NCBI using the GB2sequin webpage [29].

Genome comparison, sequence divergence, and selective pressures

Comparative chloroplast genomes of the tree *Bursera* species were performed using mVISTA online software (http://genome.lbl.gov/vista/index.shtml) [30] and the LAGAN global multiple alignment mode with *Boswellia sacra* as reference for alignment. To visualize the genes on the boundaries of the junction sites among the Cp genomes (inverted repeat regions) we used the online platform IRscope (https://irscope.shinyapps.io/irapp/) [31]. To compare divergence among the three Cp genomes, we used DnaSP v.5.1.0 [32] to set a sliding window length of 800 bp and a distance of 200 bp between loci to calculate nucleotide diversity (Pi). We examined the evolutionary rates of annotated Cp genes among the three *Bursera* species by estimating nonsynonymous (Ka) and synonymous (Ks) substitution rates and their Ka/Ks ( ) ratios using the YN model in the KaKs Calculator program v2.0 [33].

Single sequence repeat analysis

SSR markers were identified using MISA, TRF, and VMATCH, respectively, via the CPGAVAS2 web interface [27]. We specified eight repetitions for mono-, three repetitions for di-, tri-, and tetra-, and four repetitions for penta- and hexanucleotide motifs.

Phylogenetic analysis

Sequence alignment was performed using the three *Bursera* Cp genomes and seven closely related species of the Burseraceae family whose chloroplast genomes are available in NCBI GenBank

(*Commiphora gileadensis* NC 041104.1; *C. wightii* NC 036978.1; *C. foliaceae* NC 041103.1; *Garuga forrestii* OM009254.1; *G. pinnata* NC 071194.1; *G. floribunda* ON361002.1; *Bowsellia sacra* NC 029420.1) using the MAFFT software with default parameters [34]. *Rhus ovata* was used as the outgroup (NC 049061.1). We reconstructed the phylogenetic relationships for each of the four main regions of the Cp genome (LSC, SSC, IRa, and IRb) to contrast their phylogenetic relationships. The best-fitting model of sequence evolution based on BIC was performed in MEGA v.11.0.13 [35]. A Bayesian inference tree was implemented in MrBayes v3.2[36] using two independent runs, three cold chains, and one heated chain for 5,000,000 generations and sampling one tree every 1000 generations. The average standard deviation of split frequencies below 0.01 was used to evaluate convergence. The first 25% of generated trees were discarded as burn-ins and posterior probabilities were obtained from the posterior distribution of retained trees.

Results

Bursera chloroplast genome structure and content

For tree copal species, a total of 34.5 Gb of raw 151 bp Illumina reads were sequenced. The Cp genome assembly size was 159,824 bp in *B. cuneata*, 159,865 in *B. bipinnata*, and 159,872 bp in *B. palmeri.* The overall guanine and cytosine (GC) content was 38% for the three Cp genomes (Table 1).

Table 1
Summary of chloroplast genome characteristics of three *Bursera* species

Species	*Bursera bipinnata*	*Bursera cuneata*	*Bursera palmeri*
Total length (bp)	159,865	159,824	159,812
LSC length (bp)	87,656	87,625	87,668
SSC length (bp)	18,581	18,571	18,576
IRa length (bp)	26,814	26,814	26,814
IRb length (bp)	26,814	26,814	26,814
GC content (%)	38	38	38
Total gene number	135	135	135
Functional genes	90	90	90
tRNAs	37	37	37
rRNAs	8	8	8
GenBank accession	OR003920	OR003921	OR003922

The Cp genomes consisted of 135 genes, of which 90 were functional genes, 37 tRNAs, and 8 rRNAs. The three species showed the typical quadripartite structure in angiosperms and the same arrangement. An

LSC region ranges from 87,668 to 87,625 bp, two identical inverted repeats (IR) regions of 26,814, and one SSC region ranging from 18,581 to 18,571 bp. 86 loci were annotated in the LSC, 13 in the SSC, and 18 in each of the IR portions (Fig. 1). Of the annotated genes, 44 were related to photosynthesis, 24 were related to transcription and translation, 31 were RNA genes, and 7 were related to other functions (Table 2).

Table 2
List of annotated genes in chloroplast genomes of *Bursera bipinnata, B. cuneata*, and *B. palmeri.*

Gene function	Gene group	Gene name
Photosynthesis genes	Rubisco	rbcL
	Photosystem 1	psaA, psaB, psaC, psaI, psaJ
	Photosystem II	psbA, psbB, psbC, psbD, psbE, psbF, psbH, psbI, psbJ, psbK, psbL, psbM, psbT, psbZ
	ATP synthesis	atpA, atpB, atpE, atpF, atpH, atpI
	Cytochrome b/f complex	petA, petB, petD, petG, petL, petN
	Cytochrome c	ccsA
	NADPH dehydrogenase	ndhA, ndhB, ndhC, ndhD, ndhE, ndhF, ndhG, ndhH, ndhI, ndhJ, ndhK
Transcription and translation genes	Transcription	rpoA, rpoB, rpoC1, rpoC2
	Ribosomal proteins	rpl2, rpl4, rpl16, rpl20, rpl22, rpl23, rpl33, rpl32, rpl36, rps2, rps3, rps4, rps7, rps8, rps11, rps12, rps14, rps16, rps18, rps19
RNA genes	Ribosomal RNA	rrn4.5, rrn5, rrn16, rrn23
	Transfer RNA	trnA-UGC, trnC-GCA, trnD_GUC, trnE-UUC, trnF-GAA, trnG-GCC, trnG-UCC, trnH-GUG, trnI-GAU, trnK-UUU, trnL-CAA, trnL-UAG, trnL-UAA, trnM-CAU, trnN-GUU, trnP-UGG, trnQ-UUG, trnR-ACG, trnR-UCU, trnS-GGA, trnS-UGA, trnT-GGU, trnT-UGU, trnV-GAC, trnV-UAC, trnW-CCA, trnY-GUA
Other genes	RNA processing	matK
	Carbon metabolism	cemA
	Fatty acid synthesis	accD
	Proteolysis	clpP1
	Translational initiation factor	infA
	Conserved open reading frames	ycf1, ycf2
	unknown	pafI, pafII, pbf1

Comparative structure and nucleotide diversity

The structural features of the three *Bursera* species showed high conservation of coding regions relative to non-coding regions (Fig. 2). The comparative analysis of the four junction regions is shown in Fig. 3. There were a few changes in the length of the Cp genomes resulting from the expansion/contraction of IR regions. The changes were found in the JSA region with the *ndhF* gene shifted 200 and 201 bp away from the IRa-LSC junction in *B. cuneata* and *B. bipinnata* respectively, and 601 bp from the IRa border in *B. palmeri*. For the three genomes, the *rpl22* crossed the junction regions between LSC/IRb (located in JLB), and the gene *ycf1* (located in JSB) crossed the IRb/SSC junction.

The sliding window analysis comparing the three Cp genomes showed few genes with high nucleotide diversity (Pi) (i.e., mutational hotspots), being *psbJ* the gene with the highest average nucleotide diversity (0.0058), followed by *trnQ-UGG* (0.0042), *trnG-UCC* (0.0042), and then *petL* (0.0033). All of these are located in the LSC region, while no mutational hotspots were observed in other regions (Fig. 4A). For the Ka/Ks ( ) ratios high similarity among values was found for the three species, and with most genes under purifying selection, except for the genes *atpF*, *rps3*, and *ccsA* which showed a value higher than 1 suggesting positive selection (Fig. 4B).

Repeat Sequence Analysis

The identification of CpSSRs showed a very similar composition among the three Cp genomes, ranging from a total of 589 (*B. cuneata*) to 591 (*B. bipinnata*) SSRs. The most abundant SSRs for the three species were dinucleotide repeats (~ 59%), with the AT/TA motif the most frequent (~ 27%), followed by the AG/CT motif (~ 23%). Within the mononucleotide repeats, only the A/T motifs were abundant (~ 26%), while very few (< 5%) tri- and tetranucleotide motifs were present (Fig. 5).

Phylogenetic relationships

The Bayesian phylogenetic reconstructions resulted in consensus trees with very high posterior probabilities (100%). The four phylogenetic trees had similar topological structures, in which *Commiphora* was the sister taxa to *Bursera*, while *Garuga* and *Boswellia* were grouped in a different clade (Fig. 6). However, within *Bursera*, only the phylogenetic trees of the LSC and SSC regions resolved the relationships among the three species, with *B. bipinnata* and *B. palmeri* the closest species (Fig. 6a and 6b). The IRa and IRb regions did not resolve these relationships within *Bursera* (Fig. 6c and 6d).

Discussion

The three *Bursera* plastomes possess the typical quadripartite structure of the angiosperm chloroplast and are conserved in size, structure, and GC content. This suggests the high preservation of the chloroplast genome within the lineages in *Bursera*. Compared to other Cp genomes in Burseraceae species, such as *Commiphora gileadensis* (160,268 bp) and *Boswellia sacra* (160,543 bp), the Cp genome in *Bursera* was slightly smaller (158,824 to 159,872 bp), though larger than *C. wightii* (156,064 bp) [18, 19]. The average GC content in *Bursera* (38%) was almost similar to *B. sacra* (37.6%), *C. gileadensis*

(37.9%), *C. foliaceae* (37.8%) and equal to *C. wightii* (38%). The number of functional genes identified in those species (95 to 93 genes) was slightly higher than in *Bursera*, while it was lower in *C. wightii* (86 genes) [18, 19].

Protein coding genes (CDS) showed high conservatism among the three *Bursera* species, with few genes showing evidence of positive selection (*atpF*, *rps3*, and *ccsA*), and thus may have a faster evolutionary rate. The analysis of the expansion and contraction of the IR regions showed a few size variations among the three *Bursera* species and includes the location and number of base pairs in the gene *ndfh* located in the SSC-IRa junction. The gene *ycf1* crossed the IRb-SSC junction. The *rpl22* and *rps19* genes were located near the boundary of the IRa-LSC junction and with copies in the LSC-IRb junction. Compared to its closest taxa, in *Commiphora* the *ndfh* is near the SSC-IRb boundary, and the *ycf1* only crosses by a few base pairs the IRb-SSC boundary, while the *rpl22* and *rps19* were located in the same regions as in *Bursera* [18]. Moreover, within *Commiphora* species larger interspecific IR expansion and contraction changes were observed than in *Bursera*. The higher conservation of the Cp genome in *Bursera* may reflect its younger species diversification in the Middle Miocene and early Pliocene [13].

The comparison among the three *Bursera* Cp genomes revealed that mutational hotspots (high Pi diversity) were exclusively located in the LSC region with four genes, of which *psbJ* was the most diverse. A similar observation has been made for several species where high variation has been detected in genes located within the LSC region [37–39]. The four divergent genes observed in *Bursera* can be informative for further taxonomic studies aiming at resolving phylogenetic relationships within closely related species. For instance, the *psbA-trnH* and *rps16* genes used in previous phylogenetic analysis did not provide enough discriminating power between closely related species (e.g., [13, 40]) or for distinguishing the origin of putative *Bursera* hybrids [41].

Highly polymorphic markers such as SSRs are very useful in population genetic studies where the aim is to examine intraspecific genetic diversity and gene flow. As the chloroplast in most angiosperms is maternally inherited, it can provide insights into patterns of seed-mediated gene flow [42] and the role of dispersal vectors at the landscape scale [43]. Our analysis revealed that the SSRs composition was highly conserved among the three *Bursera* species. Specifically, we identified a set of ~ 590 SSRs, of which the dinucleotide repeats were the most frequent. Some of the SSRs identified could potentially be used for the development of a set of polymorphic microsatellite markers for population genetic studies, and which can be transferred among species in the *Bullockia* section.

Our phylogenetic analysis based on the Cp genomes of eleven Burseraceae species showed that *Commiphora* species are the closest to *Bursera*, which has been previously reported by Becerra et al., [44]. Both genera are within the subtribe Burserinae and are distinguished by one morphological floral character and the chromosome number [45]. Also, the phylogenetic trees recovered the close relationship of *Bowswellia* and *Garuga*, both within the subtribe Boswellinae. Due to the large variation within the SSC and LSC regions, the phylogenetic trees derived from them resolved the relationships within *Bursera*,

being *B. bipinnata* and *B. palmeri* closely related. The close relationship between these two species agrees with a previous study [13].

Conclusion

In this study, we reported and examined the first complete chloroplast genomes of the genus *Bursera*. The Cp genomes size ranged from 159,824 to 159,872 bp, GC content of 38%, and had 135 genes of which 90 were protein-coding genes, 37 tRNAs, and 8 rRNAs. Few size variations in the inverted regions were observed between the three *Bursera* species, whereas a total of 589 to 591 simple sequence repeats were identified. As *Bursera* trees are key elements of the TDFs biodiversity in the Neotropics, which have been poorly studied using genetic data, our research contributes to the development of their genomic resources for future ecological, evolutionary, and taxonomic studies.

Declarations

Availability of data and materials

The complete chloroplast sequence data that support the findings of this study have been deposited in NCBI GenBank: Bursera bipinnata: OR003920, Bursera cuneata: OR003921, and Bursera palmeri: OR003922.

Acknowledgments

We would like to thank Mayra Castro Morales for performing the DNA extractions; also, to Marisol Zurita Solís, Ingrid Lara, and Antonio González Rodriguez for quantification of DNA using Qubit; and Carlos A. Cultid-Medina for providing computer equipment to perform the genome assemblies, annotation, and data analysis.

Funding

Consejo Nacional de Humanidades, Ciencia y Tecnología (CONAHCYT) (Ciencia Básica grant number to YR: CB2016-283237).

Contributions

GPL performed data curation, analysis, and wrote the original draft; YR collected samples, designed the study, performed analysis, acquired funding, and contributed to editing the original draft. All authors reviewed the manuscript.

Competing Interest

The authors declare that they have no conflict of interest.

Ethical Approval and informed consent

This article does not contain any studies with human participants or animals performed by any of the authors.

References

1. McVaugh R, Rzedowski J (1965) Synopsis of the Genus Bursera L. in Western Mexico, with Notes on the Material of Bursera Collected by Sesse & Mocino. Kew Bull 18:317. https://doi.org/10.2307/4109252
2. Becerra JX (2003) Evolution of Mexican Bursera (Burseraceae) inferred from ITS, ETS, and 5S nuclear ribosomal DNA sequences. Mol Phylogenet Evol 26:300–309. https://doi.org/10.1016/S1055-7903(02)00256-7
3. Evans PH, Becerra JX, Venable DL, Bowers WS (2000) Chemical analysis of squirt-gun defense in Bursera and counterdefense by chrysomelid beetles. J Chem Ecol 26:745–754. https://doi.org/10.1023/A:1005436523770/METRICS
4. Rzedowski J, Medina Lemos R, Calderón de Rzedowski G (2005) Inventario del conocimiento taxonómico, así como de la diversidad y del endemismo regionales de las especies mexicanas de Bursera (Burseraceae). Acta Bot Mex 70:85–111
5. Aurora Montufar L (2004) Identidad y simbolismo del copal prehispánico y reciente. Arqueología 60–71
6. CONABIO (2008) Copales, diversidad y cultura. Mexico
7. Daly DC (1993) Studies in neotropical Burseraceae. VII. Notes on Bursera in South America. Brittonia 45:240–246
8. Rzedowski J, Guevara-Féfer F (1992) Burseraceae. Flora del Bajío y de regiones adyacentes 3:46
9. Almazán-Núñez RC, Eguiarte LE, Arizmendi M, del Corcuera C (2016) P Myiarchus flycatchers are the primary seed dispersers of Bursera longipes in a Mexican dry forest. PeerJ 2016:e2126. https://doi.org/10.7717/PEERJ.2126/TABLE-4
10. Stevenson PR, Link A, Ramírez BH (2005) Frugivory and Seed Fate in Bursera inversa (Burseraceae) at Tinigua Park, Colombia: Implications for Primate Conservation1. Biotropica 37:431–438. https://doi.org/10.1111/J.1744-7429.2005.00057.X
11. Rivas-Arancibia SP, Bello-Cervantes E, Carrillo-Ruiz H et al (2015) Variaciones de la comunidad de visitadores florales de Bursera copallifera (Burseraceae) a lo largo de un gradiente de perturbación antropogénica. Rev Mex Biodivers 86:178–187. https://doi.org/10.7550/RMB.44620
12. Becerra JX (2005) Timing the origin and expansion of the Mexican tropical dry forest. Proc Natl Acad Sci U S A 102:10919–10923. https://doi.org/10.1073/PNAS.0409127102/ASSET/B5AB354D-F64E-4B99-BD00-48B8036D9D1C/ASSETS/GRAPHIC/ZPQ0280588260004.JPEG
13. de-Nova JA, Medina R, Montero JC et al (2012) Insights into the historical construction of species-rich Mesoamerican seasonally dry tropical forests: the diversification of Bursera (Burseraceae, Sapindales). New Phytol 193:276–287. https://doi.org/10.1111/J.1469-8137.2011.03909.X

14. Becerra JX, Venable DL (1999) Nuclear ribosomal DNA phylogeny and its implications for evolutionary trends in Mexican Bursera (Burseraceae). Am J Bot 86:1047–1057. https://doi.org/10.2307/2656622
15. Rosell JA, Olson ME, Weeks A et al (2010) Diversification in species complexes: Tests of species origin and delimitation in the Bursera simaruba clade of tropical trees (Burseraceae). Mol Phylogenet Evol 57:798–811. https://doi.org/10.1016/J.YMPEV.2010.08.004
16. Weeks A, Simpson BB (2004) Molecular genetic evidence for interspecific hybridization among endemic Hispaniolan Bursera (Burseraceae). Am J Bot 91:976–984. https://doi.org/10.3732/AJB.91.6.976
17. Rico Y, Lorenzana GP, Benítez-Pineda CA, Olukolu BA (2022) Development of Genomic Resources in Mexican Bursera (Section: Bullockia: Burseraceae): Genome Assembly, Annotation, and Marker Discovery for Three Copal Species. Genes (Basel) 13. https://doi.org/10.3390/genes13101741
18. Khan A, Asaf S, Khan AL et al (2019) First complete chloroplast genomics and comparative phylogenetic analysis of Commiphora gileadensis and C. foliacea: Myrrh producing trees. PLoS ONE 14:e0208511. https://doi.org/10.1371/JOURNAL.PONE.0208511
19. Khan AL, Al-Harrasi A, Asaf S et al (2017) The First Chloroplast Genome Sequence of Boswellia sacra, a Resin-Producing Plant in Oman. PLoS ONE 12:e0169794. https://doi.org/10.1371/JOURNAL.PONE.0169794
20. Daniell H, Lin CS, Yu M, Chang WJ (2016) Chloroplast genomes: diversity, evolution, and applications in genetic engineering. Genome Biology 2016 17:1 17:1–29. https://doi.org/10.1186/S13059-016-1004-2
21. Ebert D, Peakall R (2009) Chloroplast simple sequence repeats (cpSSRs): technical resources and recommendations for expanding cpSSR discovery and applications to a wide array of plant species. Mol Ecol Resour 9:673–690. https://doi.org/10.1111/J.1755-0998.2008.02319.X
22. Figueroa-Suárez MZ, González Christen J, Cardoso-Taketa AT et al (2019) Anti-inflammatory and antihistaminic activity of triterpenoids isolated from Bursera cuneata (Schldl.) Engl. J Ethnopharmacol 238:111786. https://doi.org/10.1016/J.JEP.2019.03.013
23. Healey A, Furtado A, Cooper T, Henry RJ (2014) Protocol: A simple method for extracting next-generation sequencing quality genomic DNA from recalcitrant plant species. Plant Methods 10:1–8. https://doi.org/10.1186/1746-4811-10-21/COMMENTS
24. Chen S, Zhou Y, Chen Y, Gu J (2018) fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34:i884–i890. https://doi.org/10.1093/BIOINFORMATICS/BTY560
25. Jin JJ, Yu W, Bin, Yang JB et al (2020) GetOrganelle: A fast and versatile toolkit for accurate de novo assembly of organelle genomes. Genome Biol 21:1–31. https://doi.org/10.1186/S13059-020-02154-5/FIGURES/6
26. Okonechnikov K, Golosova O, Fursov M et al (2012) Unipro UGENE: a unified bioinformatics toolkit. Bioinformatics 28:1166–1167. https://doi.org/10.1093/BIOINFORMATICS/BTS091

27. Shi L, Chen H, Jiang M et al (2019) CPGAVAS2, an integrated plastome sequence annotator and analyzer. Nucleic Acids Res 47:W65–W73. https://doi.org/10.1093/NAR/GKZ345
28. Zheng S, Poczai P, Hyvönen J et al (2020) Chloroplot: An Online Program for the Versatile Plotting of Organelle Genomes. Front Genet 11:576124. https://doi.org/10.3389/FGENE.2020.576124/BIBTEX
29. Lehwark P, Greiner S (2019) GB2sequin - A file converter preparing custom GenBank files for database submission. Genomics 111:759–761. https://doi.org/10.1016/J.YGENO.2018.05.003
30. Frazer KA, Pachter L, Poliakov A et al (2004) VISTA: Computational tools for comparative genomics. Nucleic Acids Res 32. https://doi.org/10.1093/NAR/GKH458
31. Amiryousefi A, Hyvönen J, Poczai P (2018) IRscope: an online program to visualize the junction sites of chloroplast genomes. Bioinformatics 34:3030–3031. https://doi.org/10.1093/BIOINFORMATICS/BTY220
32. Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25:1451–1452. https://doi.org/10.1093/BIOINFORMATICS/BTP187
33. Wang D, Zhang Y, Zhang Z et al (2010) KaKs_Calculator 2.0: A Toolkit Incorporating Gamma-Series Methods and Sliding Window Strategies. Genomics Proteom Bioinf 8:77. https://doi.org/10.1016/S1672-0229(10)60008-3
34. Katoh K, Standley DM (2013) MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. Mol Biol Evol 30:772–780. https://doi.org/10.1093/MOLBEV/MST010
35. Tamura K, Stecher G, Kumar S (2021) MEGA11: Molecular Evolutionary Genetics Analysis Version 11. Mol Biol Evol 38:3022–3027. https://doi.org/10.1093/MOLBEV/MSAB120
36. Ronquist F, Teslenko M, van der Mark P et al (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst Biol 61:539–542. https://doi.org/10.1093/sysbio/sys029
37. Cheng H, Li J, Zhang H et al (2017) The complete chloroplast genome sequence of strawberry (Fragaria × ananassa Duch.) and comparison with related species of Rosaceae. PeerJ 2017:e3919. https://doi.org/10.7717/PEERJ.3919/SUPP-7
38. Li C, Cai C, Tao Y et al (2021) Variation and Evolution of the Whole Chloroplast Genomes of Fragaria spp. (Rosaceae). Front Plant Sci 12:2293. https://doi.org/10.3389/FPLS.2021.754209/BIBTEX
39. Tyagi S, Jung JA, Sun Kim J, Youn Won S (2020) Comparative Analysis of the Complete Chloroplast Genome of Mainland Aster spathulifolius and Other Aster Species. Plants (Basel) 9. https://doi.org/10.3390/PLANTS9050568
40. Weeks A, Simpson BB (2007) Molecular phylogenetic analysis of Commiphora (Burseraceae) yields insight on the evolution and historical biogeography of an impossible genus. Mol Phylogenet Evol 42:62–79. https://doi.org/10.1016/j.ympev.2006.06.015
41. Melecio EQ, Rico Y, Noriega AL, Rodríguez AG (2021) Molecular evidence and ecological niche modeling reveal an extensive hybrid zone among three Bursera species (section Bullockia). PLoS ONE 16:e0260382. https://doi.org/10.1371/JOURNAL.PONE.0260382

42. Jansen RK, Cai Z, Raubeson LA et al (2007) Analysis of 81 genes from 64 plastid genomes resolves relationships in angiosperms and identifies genome-scale evolutionary patterns. Proc Natl Acad Sci U S A 104:19369–19374. https://doi.org/10.1073/PNAS.0709121104/SUPPL_FILE/09121FIG9.PDF
43. Auffret AG, Schmucki R, Reimark J, Cousins SAO (2012) Grazing networks provide useful functional connectivity for plants in fragmented systems. J Veg Sci 23:970–977. https://doi.org/10.1111/j.1654-1103.2012.01413.x
44. Becerra J, Daly DC, Simpson BB (2005) The phylogenetic history and biogeography of the frankincense and myrrh family (Burseraceae) based on nuclear and chloroplast sequence data. Mol Phylogenet Evol 35:85–101
45. Bawa K (1973) Chromosome numbers of tree species of a lowland tropical community. J Arnold Arboretum 54:422–434

Figures

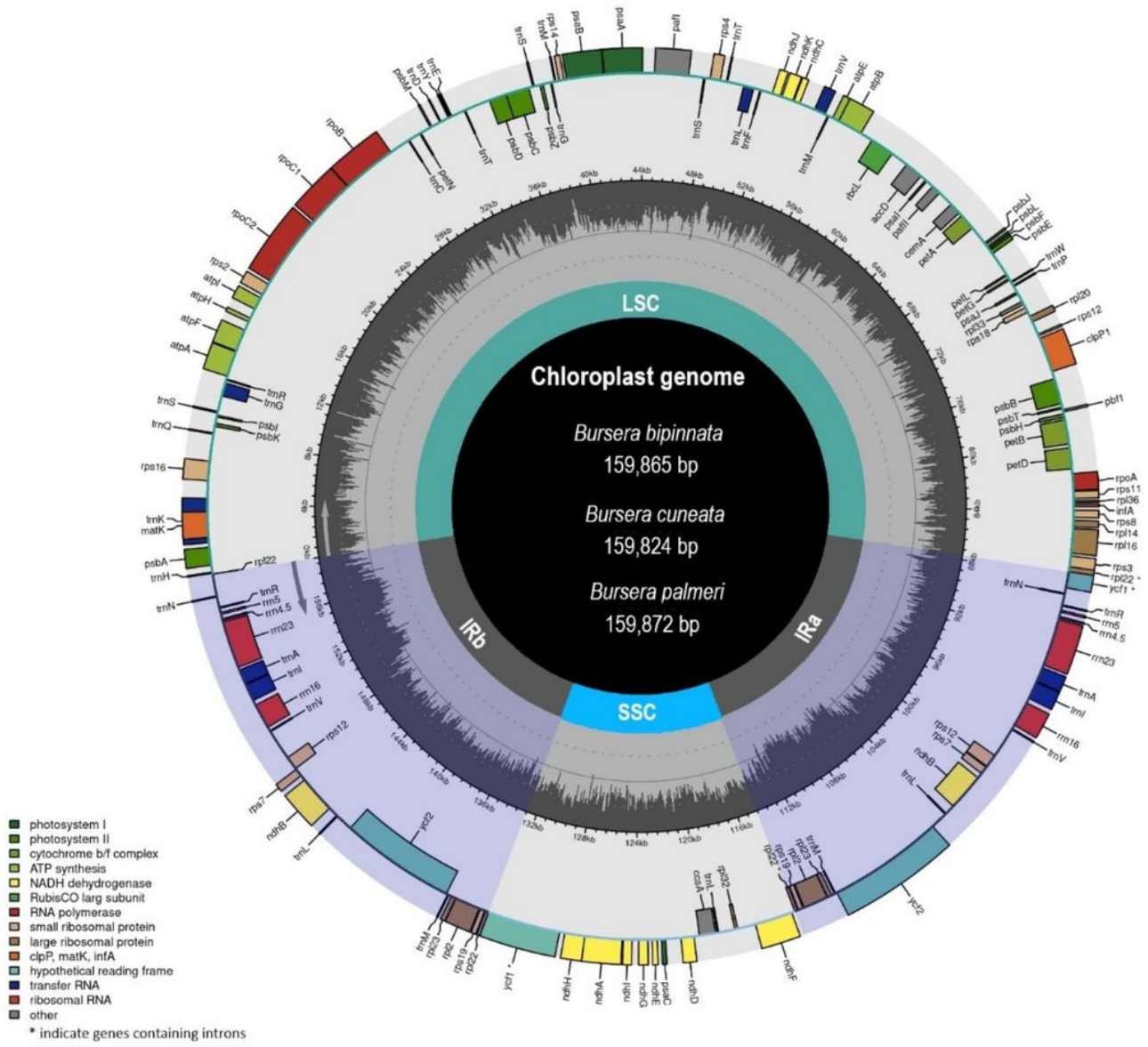

Figure 1

The circular map of the *Bursera* chloroplast genome structure. Genes are shown along the inner and outer sides of the circle; genes outside the circle are transcribed counterclockwise and those in the inside are transcribed clockwise. Genes with different functions are displayed in different colors.

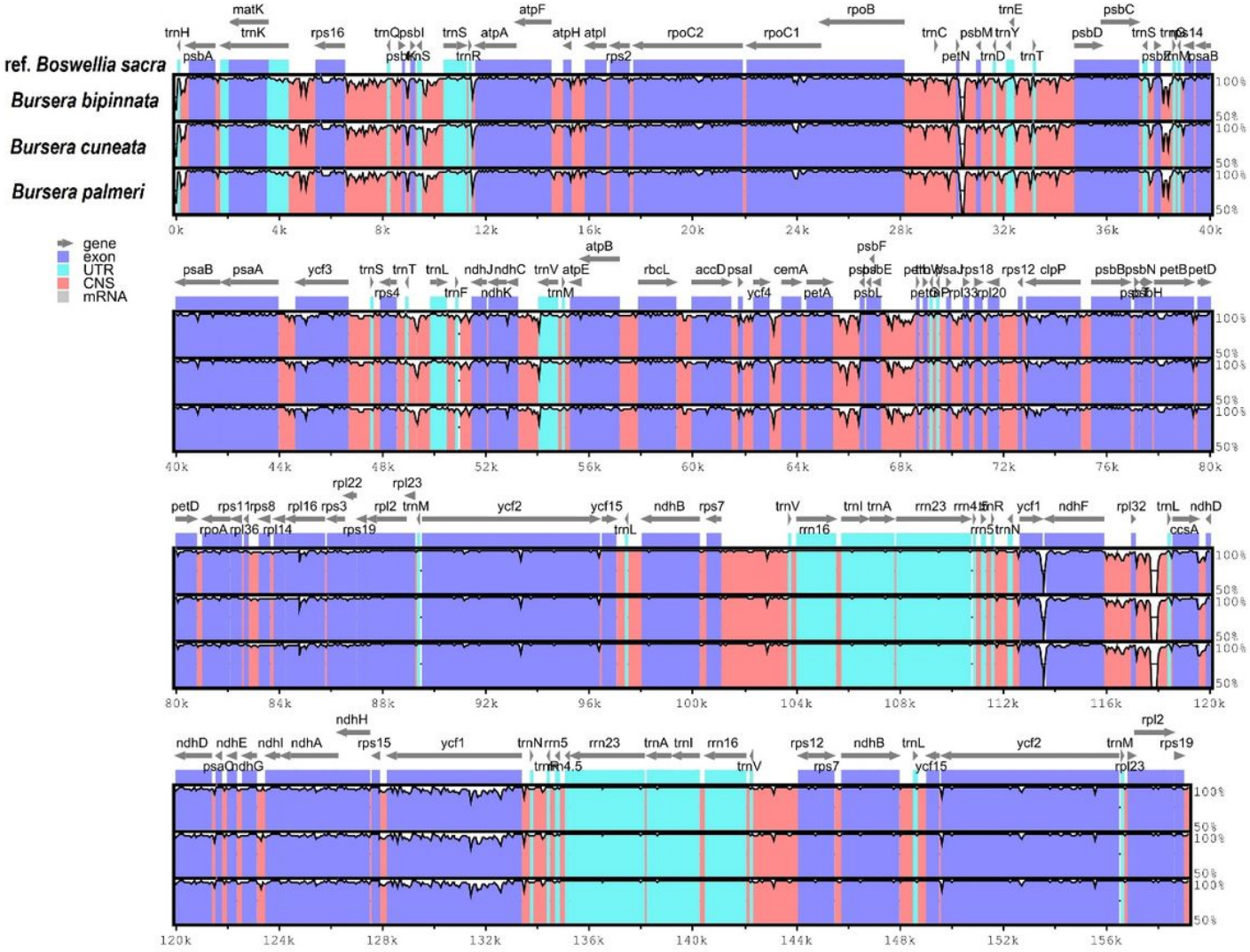

Figure 2

Visualization of the chloroplast genome alignment of three *Bursera* species using *Boswellia sacra* as a reference in mVISTA software. The vertical line indicates the percentage of identity; coding regions are marked in purple and non-coding in pink.

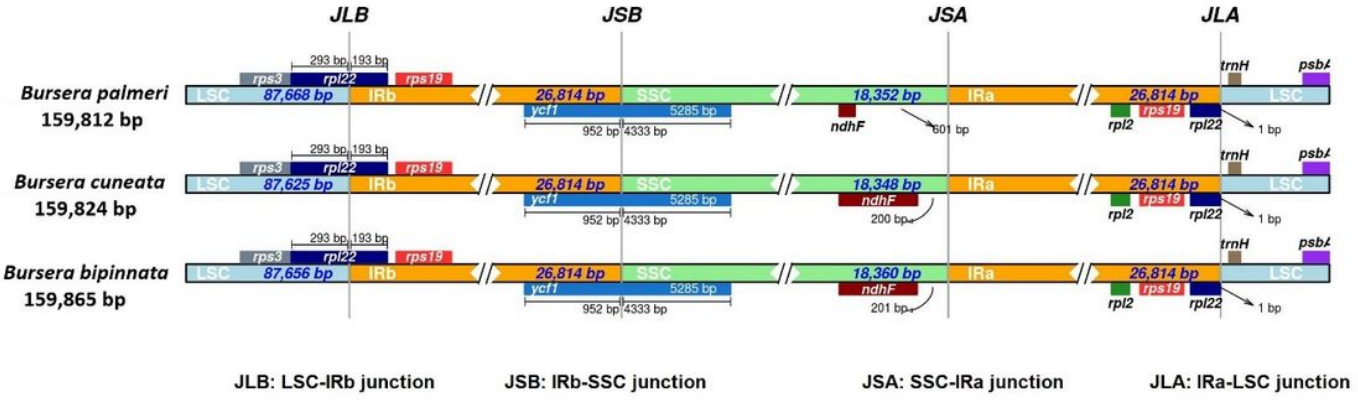

Figure 3

Comparison of the junction regions, JLB, JSB, JSA, and JLA, of the chloroplast genomes of three *Bursera* species. Different genes are displayed by colored boxes and the gaps between the genes and the boundaries are denoted by the base lengths (bp).

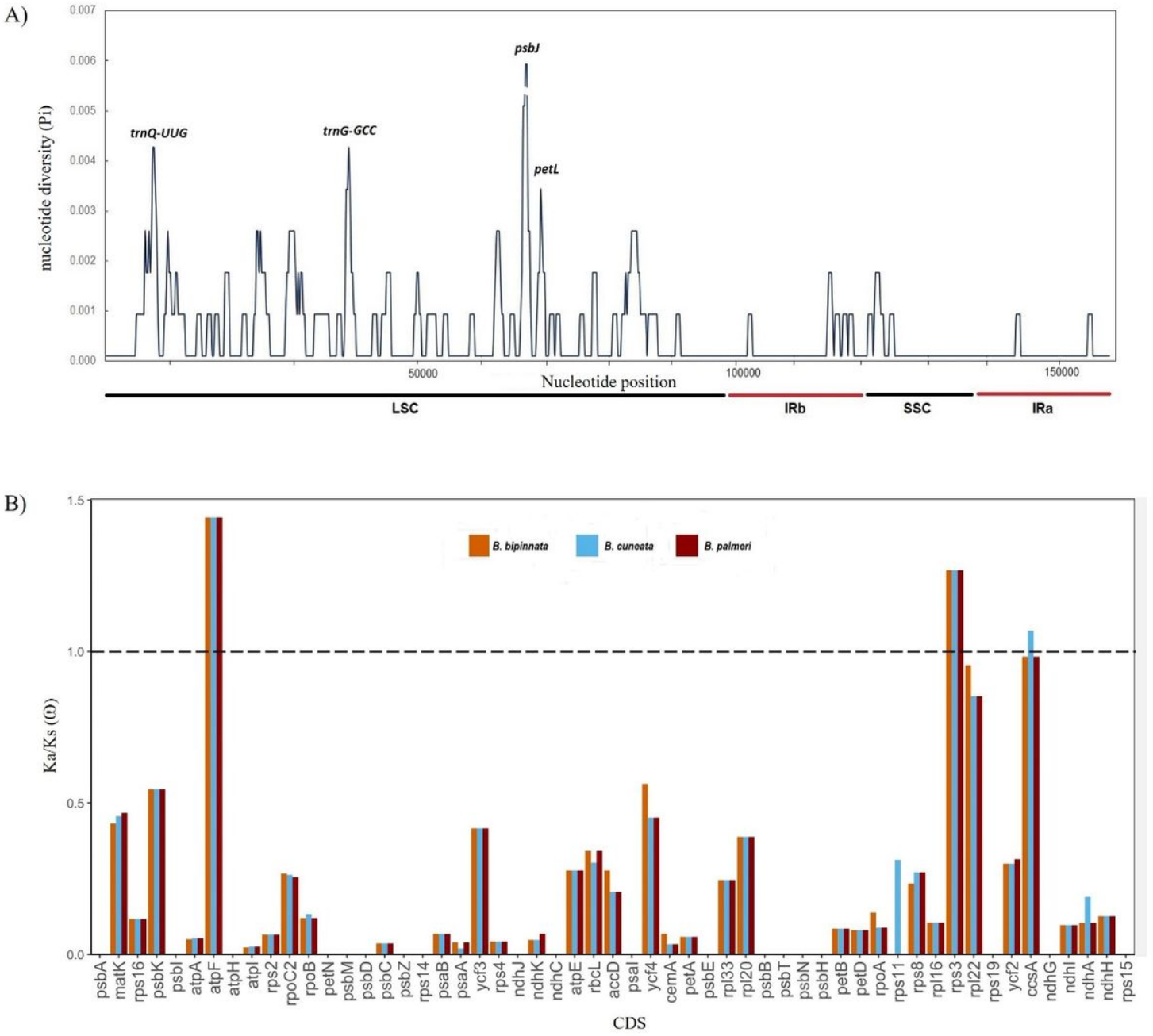

Figure 4

(A) Sliding window analysis of three chloroplast genomes of *Bursera*section *Bullockia*. The x-axis denotes the nucleotide position along the sequences and the y-axis the nucleotide diversity value (Pi) per window. (B) Analysis of Ka/Ks (w) ratios for 55 protein-coding genes (CDS) among the chloroplast genomes of three *Bursera* species.

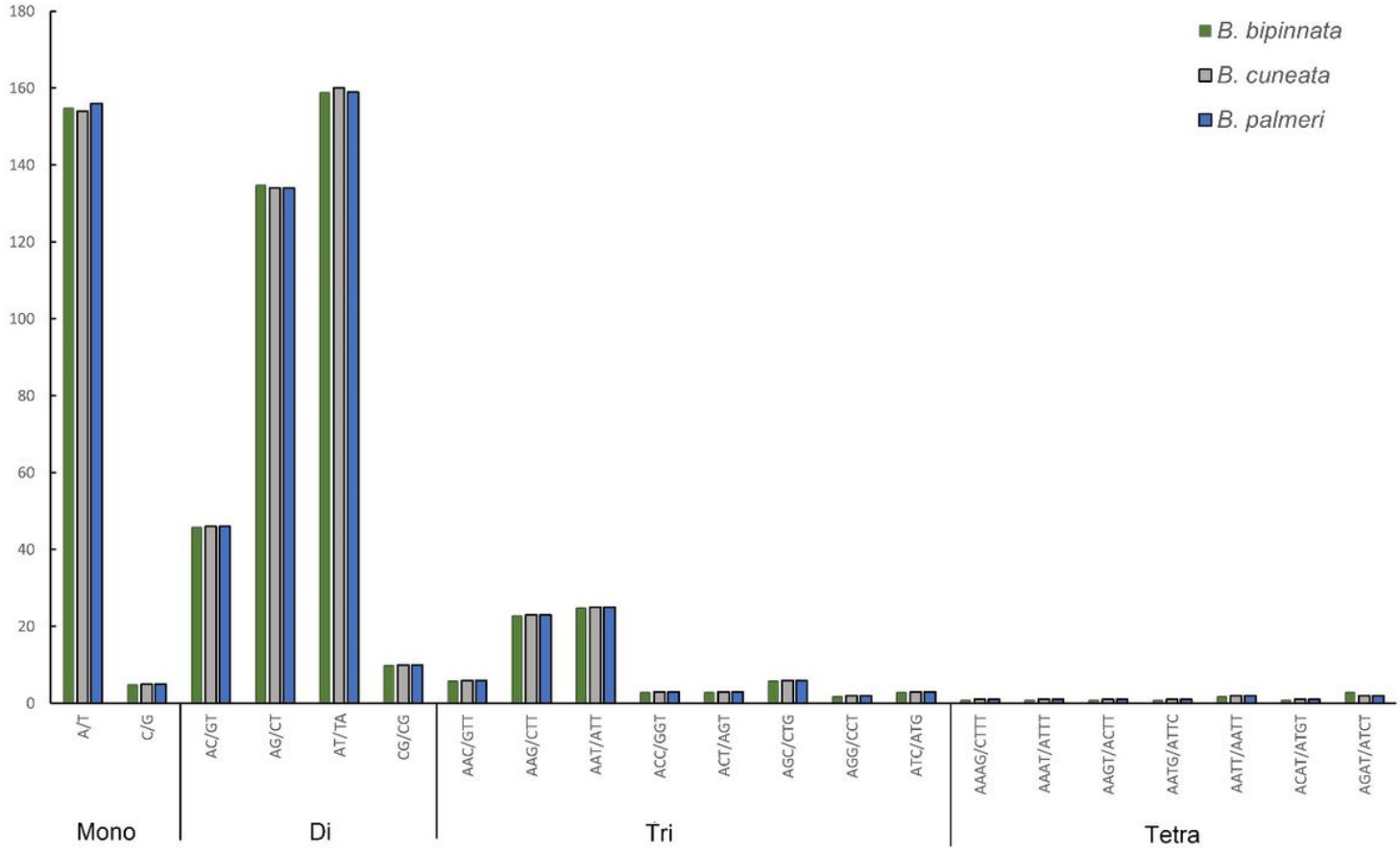

Figure 5

Comparison of the frequency distribution of the types of simple sequence repeat markers (SSR) present in the chloroplast genomes of three *Bursera*species.

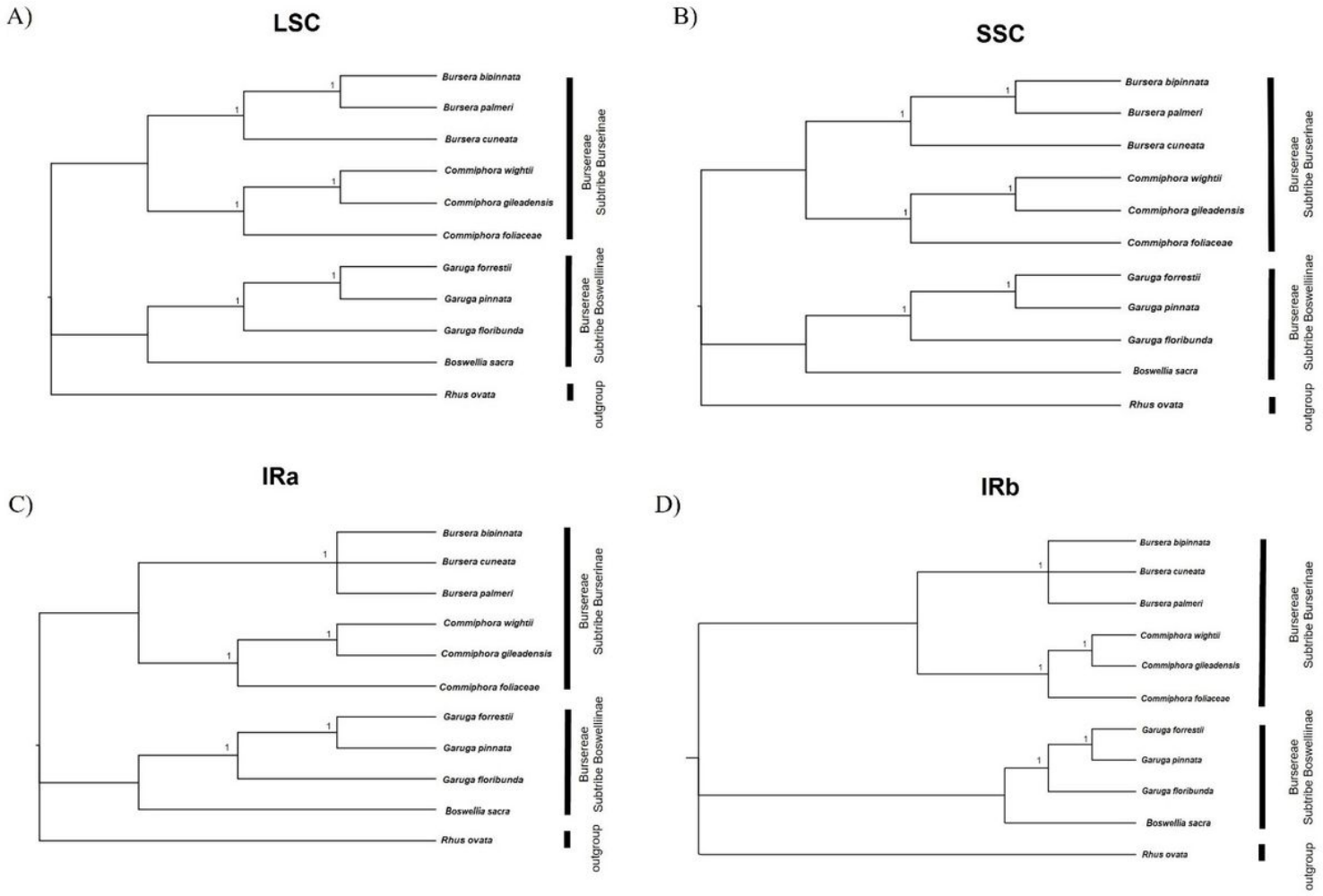

Figure 6

Bayesian phylogenetic trees of eleven species of Burseraceae per chloroplast genome region: (A) Long (LSC) and (B) short single-copy (SSC), and the two inverted repeats (C) (IRa) and (D) (IRb). Values besides the branch represent the posterior probabilities.